

1/11

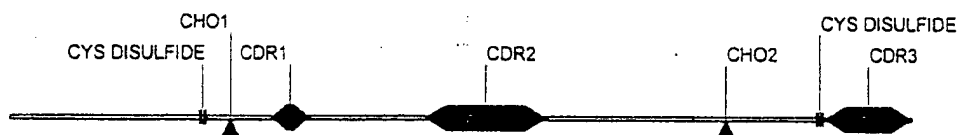
Figure 1. Alignment of Thy-1 and 8E5 VH

		1		50
EcFv-1.15* 8E5 VH	(1)	-----QLOQS--G--AELVKPGA	SVKMS	MAAGYTFFNYWM
Thy1 human	(1)	MNLAISIALLLTVLQVSRGQKVTSLTACLVDQSLRLD	CRHENTSSSPIQY	
Consensus	(1)	QL	SLKL CK	S S
		51		100
EcFv-1.15* 8E5 VH	(33)	HWVKQSPGQGLEWIGTIDPADSYTSYNQNF	-----	KDKA
Thy1 human	(51)	ESLIRETKKKHVLFGTIDGVPDHTYRSRTN	ETSKYHMKVLYLSAFTSK	DEG
Consensus	(51)	F R GTI D NF		KD A
		101		150
EcFv-1.15* 8E5 VH	(67)	FLVDPKPSSTAYMQESSLTFGDSAVYFCARE	SYRYFFDY	NGHGTTLTV
Thy1 human	(101)	TYTCALHHSGHSPPISSQNVTVLRDKLVKCE	ISLLAQNTSWLLLLLESL	
Consensus	(101)	T T S ISS	EGI	W LSL
		151	161	
EcFv-1.15* 8E5 VH	(117)	SSAKTTPKE--		
Thy1 human	(151)	SLLQATDFMSL		
Consensus	(151)	S T L		

Figure 1B. Design of a single Ig domain CDR binding polypeptide based on the Thy-1 structure.

Qvsrgqkvtsltacldvqslrldcrhentsssnym
Hfsltretkkhvlfgtidpadsytsynqnfkdegtytc
Alhhsghspissqnvtvlrdklvkcegyyryyfdy

Figure 1C. Diagram of a single Ig domain CDR carrier based on the Thy-1 structure.



Thy1/8E4 VH synthetic CDR binding polypeptide
111 aa

2/11

FIGURE 2A

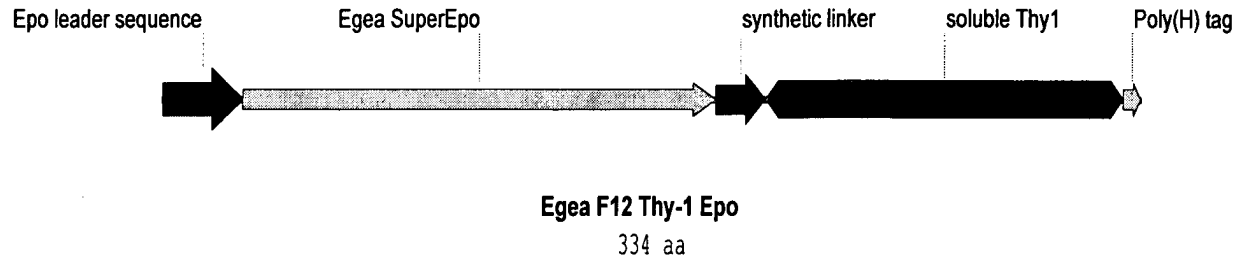


FIGURE 2B

1	MGVHECPAWL	WLLLSLLSLP	LGLPVLGAPP	RLICDSRVLE	RHLLEAKEAE
51	SITTGCVEDC	SLNENITVPD	SKVNFYAWKR	MEVGQQAVEV	WQGLALLSEA
101	VLRGQALLVI	SSQPWEPLQL	HVDKAVSGLR	SLTTLLRALG	AQKEAISPPD
151	AASAAPLRTI	TADTFRKLFR	VYPNFLRGKL	KFYTGACRG	GGGSGGGGE
201	FGGGGSQKVT	SLTACLVDQS	LRLDCRHENT	SSSPIQYEFS	LTRETKKHVL
251	FGTVGVPEHT	YRSRTNFTSK	YHMKVLYLSA	FTSKDEGTYT	CALHHSQHSP
301	PISSQNVTVL	RDKLVKCEGI	SLLAQNTSHH	HHHH	

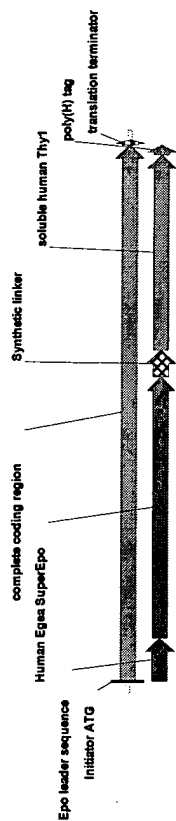
3/11

FIGURE 2C

```

HindIII
~~~~~
      M G V H E C P A W L W L L L S L L S L P
1  GATTGGCGAA GCTTGGAGGA ATGGGCGTGC ACGAGTGTCC CGCCTGGCTG TGGCTGCTGC TGAGCCTGCT GAGCCTGTCCC
   L G L P V L G A P P R L I C D S R V L E R H L L E A K .
81 CTGGGCGTGC CCGTGTGGG CGCCCCCCCC CGGCTGATCT GCGACAGCCG GGTGCTGGAG CGGCACCTGC TGGAGGCCAA
   . E A E S I T T G C V E D C S L N E N I T V P D S K V N .
161 GGAGGCCGAG AGCATCACCA CCGGCTGCGT GGAGGACTGC AGCCTGAACG AGAACATCAC CGTGCCCGAC AGCAAGGTGA
   . F Y A W K R M E V G Q Q A V E V W Q G L A L L S E A
241 ACTTCTACGC CTGGAAGCGG ATGGAGGTGG GCCAGCAGGC CGTGGAGGTG TGGCAGGGCC TGGCCCTGCT GAGCGAGGCC
   V L R G Q A L L V I S S Q P W E P L Q L H V D K A V S .
321 GTGCTGCGGG GCGAGGCCCT GCTGGTGATC AGCAGCCAGC CCTGGGAGCC CCTGCAGCTG CACGTGGACA AGGCCGTGAG
   . G L R S L T T L L R A L G A Q K E A I S P P D A A S A .
401 CCGCCTGCGG AGCCTGACCA CCCTGCTGCG GGCCCTGGGC GCCCAGAAGG AGGCCATCAG CCCCCCGAC GCCGCCAGCG
   . A P L R T I T A D T F R K L F R V Y P N F L R G K L
481 CCGCCCCCT GCGGACCATC ACCGCCGACA CTTCCGGAA GCTGTCCGG GTGTACCCCA ACTTCCTGCG GGGCAAGCTG
   K F Y T G E A C R G G G G G S G G G G E F G G G G S Q .
561 AAGTTCTACA CCGCGAGGC CTGCCGGGGC GCGCGCGCGC GCAGCGGCGG CGGCAGCGAG TTCGGCGGCG GCGGCAGCCA
   . K V T S L T A C L V D Q S L R L D C R H E N T S S S P .
641 GAAGGTGACC AGCCTGACCG CCTGCCTGGT GGACCAGAGC CTGCGGCTGG ACTGCCGCA CGAGAACACC AGCAGCAGCC
   . I Q Y E F S L T R E T K K H V L F G T V G V P E H T
721 CCATCCAGTA CGAGTTCAGC CTGACCCGGG AGACCAAGAA GCACGTGCTG TTCGGCACCG TGGGCGTGCC CGAGCACACC
   Y R S R T N F T S K Y H M K V L Y L S A F T S K D E G .
801 TACCGGAGCC GGACCAACTT CACCAGCAAG TACCACATGA AGGTGCTGTA CCTGAGCGCC TTCACCAGCA AGGACGAGGG
   . T Y T C A L H H S G H S P P I S S Q N V T V L R D K L .
881 CACCTACACC TGCGCCCTGC ACCACAGCGG CCACAGCCCC CCCATCAGCA GCCAGAACGT GACCGTGCTG CGGGACAAGC
                                     BamHI
                                     ~~~~~
   . V K C E G I S L L A Q N T S H H H H H H * * *
961 TGGTGAAGTG CGAGGGCATC AGCCTGTGG CCCAGAACAC CAGCCACCAC CACCACCACC ACTGATGATA AGATCGGATC
   BamHI
   ~
1041 CTAGGCTTCC
```

4/11



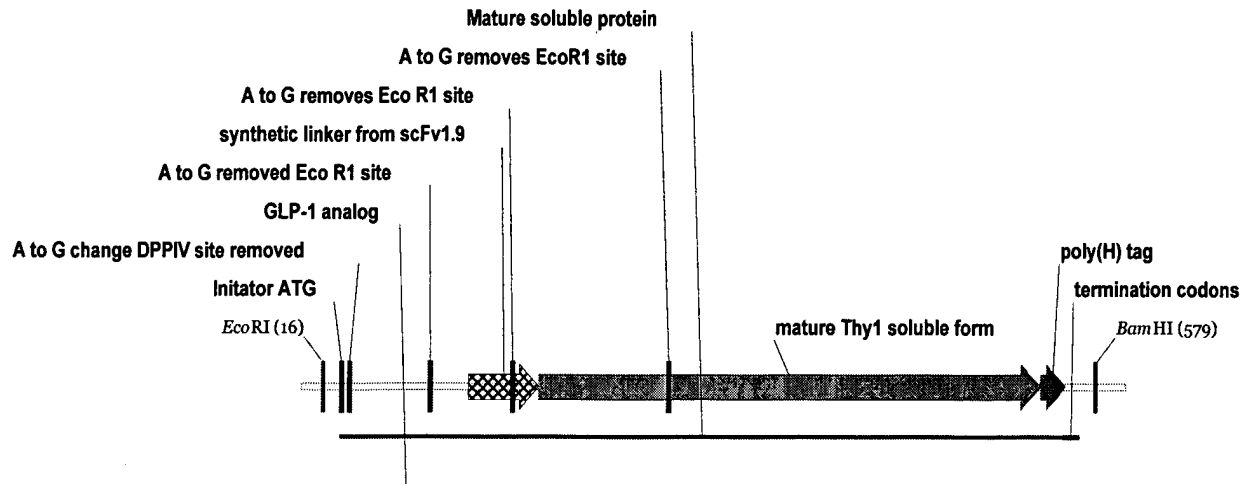
1 GATTGGCGAA GCTTGGAGGA ATGGCGTGC ACAGTGGCCC CGCCTGGCTG TGGCTGCTGC TGAGCTGCTT GAGCTGGCCC CTGGGCTTGC CCGTGTCTGG
CTAACCCGCTT CGAACCCTCT TACCCGCACG TGCTCACGGG GCGACCCGAC ACCGACGACG ACTCGGACGA CTGCGACGGG GACCCGGAGC GGCACGACCC
101 GlyProPro ArgLeuileCys AspSerArg ValLeuGlu ArgHisLeuLeu GluAlaGlu SerIleThrThr GlyCysVal GluAspCys
CGCCCCCCCC CGGCTGATCT GCGACAGCCG GGTGCTGGAG CGGACCTGC TGGAGGCCAA GGAGGCCGAG AGCATCACCA CCGCTGCGT GGAGGACTGC
GCGGGGGGGG CGGACTAGA CGCTGTCGGC CCAGACCTC GCGTGGAGC ACCTCCGTT CCGTGGCTC TCGTAGTGTG GCGCGACGCA CTTCTGACG
SerLeuAsnGlu AsnIleThr ValProAsp SerLysValAsn PheTyrAla TrpLysArg MetGluValGly GlnGlnAla ValGluVal TrpGlnGlyLeu
201 AGCTGAACG AGAATATCAC CGTGCCCGAC AGCAAGTGA ACTTCTAGC CTGGAAGCGG ATGAGGTGG GCGACAGGC CTGGAGGTG TGGCAGGGCC
TCGACTTGC TCTTGTAGT GCACGGGCTG TCGTTCACCT TGAAGATGG GACCTTCGCC TACTCCACC CGTCTGTCG GCACCTCCAC ACCGTCCCGG
LalaLeuLeu SerGluAla ValLeuArgGly GlnAlaLeu LeuValIle SerSerGlnPro TrpGluPro LeuGlnLeu HisValAspLys AlaValSer
301 TGGCCTGTCT GAGCAGAGCC GTGTGCGGG GCCAGGCCCT GCTGTGATC AGCAGCCAGC CTTGGAGGC CTTGACGCTG CACGTGGACA AGGCCGTGAG
ACCGGGACGA CTCGTCCGG CACGAGCCC CGGTCCGGGA GCACACTAG TCGTGGTGC GGACCTCGG GGACGTCGAC GTGCACCTGT TCCGCGACTC
GlyLeuArg SerLeuThrThr LeuLeuArg AlaLeuGly AlaLeuSer ProProAsp AlaAlaSerAla AlaProLeu ArgThrIle
401 CGGCTGTGG AGCTGTACCA CCTGTCTGG GGCCTGGG GCGGAGCCG CGGTCTTCC TCCGTTAGT GGGGGGGCTG CCGCGGTGC GCGGGGGGA CGCTGTGTAG
GCCGAGCC TCGACTGT GGGACGACGC CCGGAGCCG CGGTCTTCC TCCGTTAGT GGGGGGGCTG CCGCGGTGC GCGGGGGGA CGCTGTGTAG
ThrAlaAspThr PheArgLys LeuPheArg ValTyrProAsn PheLeuArg GlyLysLeu LysPheTyrThr GlyGluAla CysArgGly GlyGlyGlyGly
501 ACCCGGACA CTTCCGGAA GCTGTCCGG GTGTACCCCA ACTTCTGCG GGGCAAGCTG AGTTCTACA CCGCGAGGC CTGCGGGGGC GCGCGGGCG
TGGCGCTGT GGAAGGCTT GCACAAGGC CACATGGGT TGAAGGAGC CCGCTTCGAC TTCAAGATGT GGCCTCTCC GACGCGCCG CCGCGCGCG
GserGlyGly GlyGlyGlu PheGlyGlyGly GlySerGln LysValThr SerLeuThrAla CysLeuVal AspGlnSer LeuArgLeuAsp CysArgHis
601 GCAGCGCGG CCGCGCGAG TCGCGCGG CCGCGAGCCA GAAGTGACC AGCTGACC CCGCTGCTG TCGGACTGG GACCGGACCA CTTGTCTCG GACCGCGAC TGACGGCCGT
CGTCCCGC GCCCGCGCTC AAGCGCGCTC CCGCTGCTG CTTCCACTGG TCGGACTGG GACCGGACCA CTTGTCTCG GACCGCGAC TGACGGCCGT
GluAsnThr SerSerSerPro IleGlnTyr GluPheSer LeuThrArgGlu ThrLysLys HisValLeu PheGlyThrVal GlyValPro GluHisThr
701 CGAGAACC AGCAGCAGC CCATCCAGTA CGAGTTCAG CTGACCCGG AGACCAAGAA GCAGTGTCTG TTCGGCACCG TGGGCGTGC CGAGCACAC
GCTCTTGTG TCGTGTCTCG GGTAGTCTAT GCTCAAGTC GACTGGGCC TCTGTCTCT CGTCACGAC AACCGTGGC ACCCGCACCG GCTCGTGTG
TyrArgSerArg ThrAsnPhe ThrSerLys TyrHisMetLys ValLeuTyr LeuSerAla PheThrSerLys AspGluGly ThrTyrThr CysAlaLeuHis
801 TACCGAGCC GGACCAACTT CACCAAGAG TACCATGTA AGGTGCTGTA CCGTGGCCG AAGTGTCTCT TCCTGTCTCC CACTACAC TGCGCCCTGC
ATGCTCTCG CTTGTTGAA GTGTCTGTTT ATGTGTACT TCCAGACAT GACTGTGCTG GAGTGTGCTG TCGTGTCTCT TCCTGTCTCC CACTACAC TGCGCCCTGC
HHISerGly HisSerPro ProIleSerSer GlnAsnVal ThrValLeu ArgAspLysLeu ValLysCys GlnGlyIle SerLeuLeuAla GlnAsnThr
901 ACCACAGCG CCACAGCCC CCGATCAGCA GCCAGAACGT GACCGTGTG CCGGACAAG TGGTGAAGTG CGAGGGCATC AGCTGTCTG CCGAGAACAC
TGGTGTGCGC GTGTGCGGG GGTGTGCTG CCGTCTTGCA CTGGACACGAC GCCCTGTTCC ACCACTTCAC GCTCCCGTAG TCGGACGACC GGGTCTTGTG
SerHisHis HisHisHisHis
1001 CAGCCACCAC CACCAACCAC ACTGATGATA AGATCGGATC CTAGGCTTCC
GTGCGTGGTG GTGCGTGGTG TGACTACTAT TCTAGCCTAG GATCCGAAGG

FIGURE 3

5/11

FIGURE 4A

Egea Thy-1_Glucagon-Like Peptide 1 Non-Immunoglobulin Carrier
Polypeptide



EgeaA42 GLP Thy1 Carrier
600 bp

FIGURE 4B

176 aa

```
1  MHGEGTFTSD VSSYLEGQAA KEFIAWLVKG RGGGGGSGGG GEFGGGGSQK
51  VTSLTACLVD QSLRLDCRHE NTSSSPIQYE FSLTRETKKH VLFGTGVGPE
101 HTYRSRTNFT SKYHMKVLYL SAFTSKDEGT YTCALHHS GH SPPISSQNVT
151 VLRDKLVKCE GISLLAQNTS HHHHHH
```

6/11

FIGURE 4C

```

                                EcoRI
                                ~~~~~
                                M  H  G  E  G  T  F  T  S  D  V  S  S  Y  L  E  G  Q  .
1   AGTCCGGGAT TTAAGAATTC AGCTGTCCAT GCACGGTGAA GGTACCTTCA CCTCTGACGT TTCTTCTTAC CTGGAAGGTC
   . A A K E F I A W L V K G R G G G G G S G G G G E F G
81  AGGCGGCGAA AGAGTTCATC GCGTGGCTGG TTAAAGGTCG TGGTGGTGGT GGTGGTTCTG GTGGTGGTGG TGAGTTCGGT
   G G G S Q K V T S L T A C L V D Q S L R L D C R H E N .
161 GGTGGTGGTT CTCAGAAAGT TACCTCTCTG ACCGCGTGCC TGGTTGACCA GTCTCTGCGT CTGGACTGCC GTCACGAAAA
   . T S S S P I Q Y E F S L T R E T K K H V L F G T V G V .
241 CACCTCTTCT TCTCCGATCC AGTACGAGTT CTCTCTGACC CGTGAAACCA AAAACACGT TCTGTTGCGT ACCGTTGGTG
   . P E H T Y R S R T N F T S K Y H M K V L Y L S A F T
321 TTCCGGAACA CACCTACCGT TCTCGTACCA ACTTCACCTC TAAATACCAC ATGAAAGTTC TGTACCTGTC TGCCTTCACC
   S K D E G T Y T C A L H H S G H S P P I S S Q N V T V .
401 TCTAAAGACG AAGGTACCTA CACCTGCGCG CTGCACCACT CTGGTCACTC TCCGCCGATC TCTTCTCAGA ACGTTACCGT
   . L R D K L V K C E G I S L L A Q N T S H H H H H H * * .
481 TCTGCGTGAC AAATGGTTA AATGCGAAGG TATCTCTCTG CTGGCGCAGA ACACCTCTCA CCACCACCAC CACCACTGAT
                                BamHI
                                ~~~~~
   . *
561 AATGAGATCT TGAGGCCGGA TCCGCTTAAG ATCCCGGCAA
```

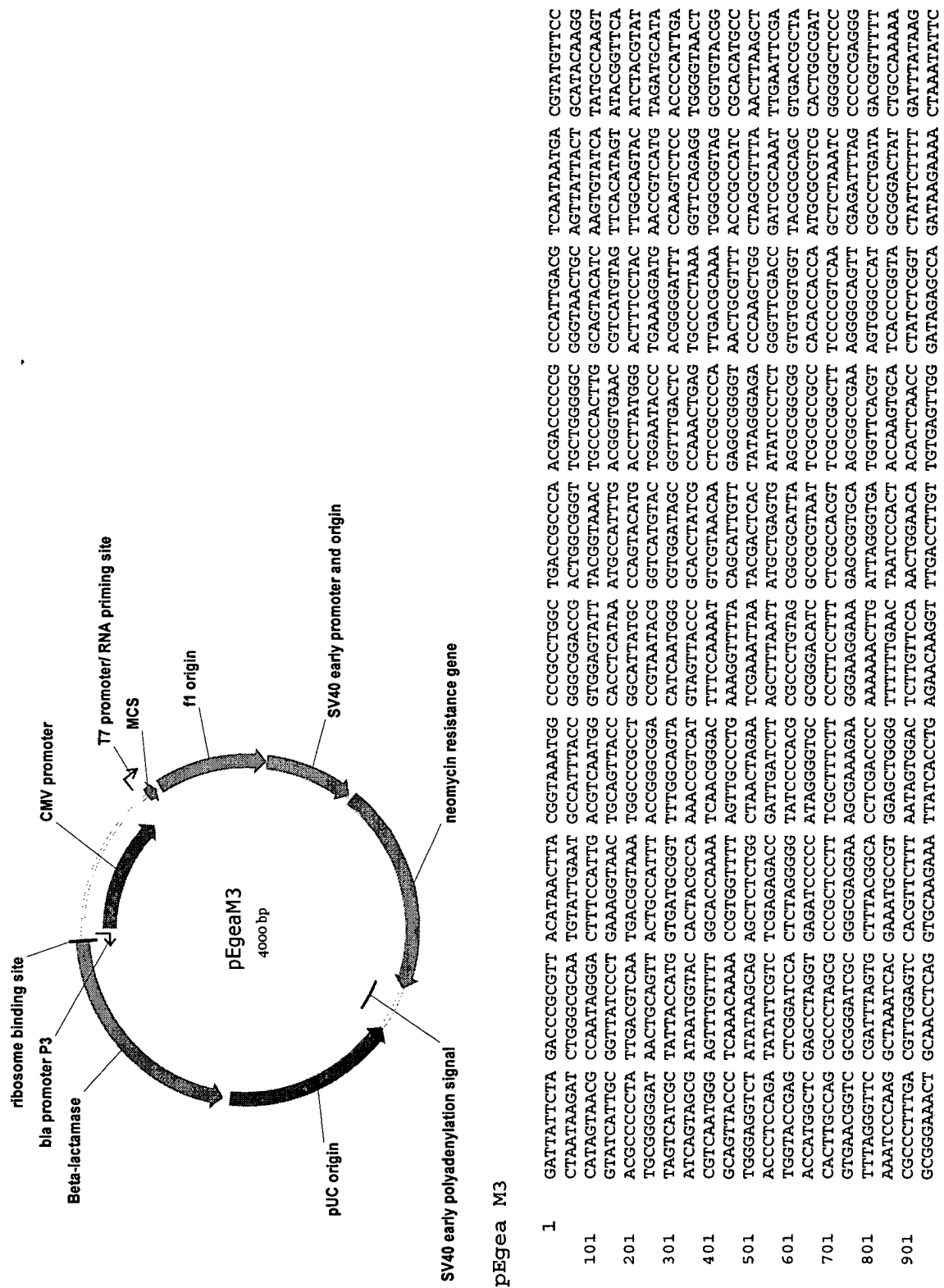


FIGURE 5

8/11

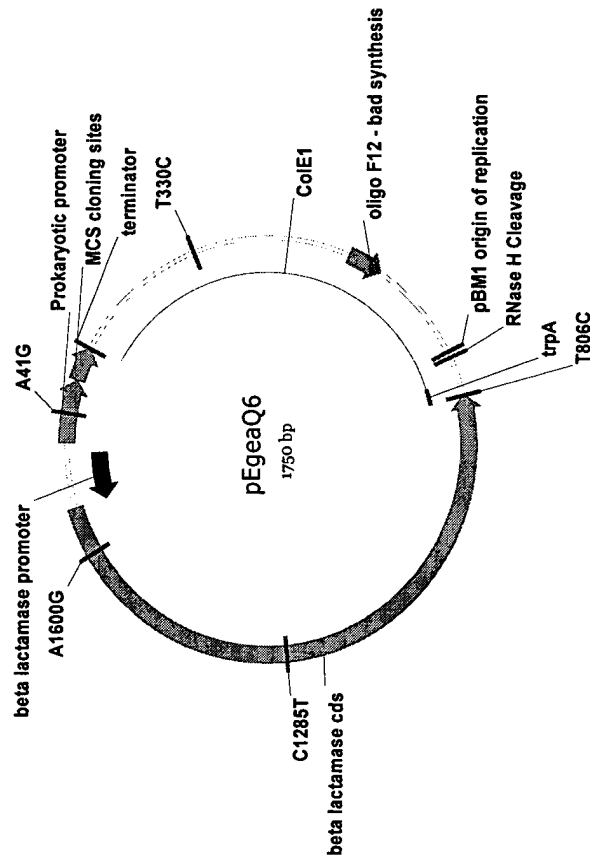
1001 GGATTCTGCC GATTTCGGCC TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAAGC CGAATTAATTT CTGTGGAATG TGTGTCAGTT AGGTGTGGA
1101 CCTAAACGG CTAAGCCCGG ATAAACAAAT TTTTACTCGA CTAATATGTT TTTAAATTGC GCTTAATTAAG GACACTTTAC ACACAGTCAA TCCACACCT
1201 AAGTCCCCAG GCTCCCCCAG AGGCAAGAAGT ATGCAAGCA TGCACTCTCA TTAGTCAGCA ACCAGGTGTG GAAAGTCCCC AGGTCCCCA GCAGGAGAA
1301 TTAGGGGTC CGAGGGTGC TCCGTCTTCA TACGTTTCGT ACCTAGAGTT AATCAGTCTG TGTGTCACAC CTTTCAGGGG TCCAGAGGGT CGTCCGTCTT
1401 GTATGCAAGC CATGCATCTC AATAGTTCAG CAACCATATG CCGCCCCCTA ACTCCGCCCA TCCCGCCCTT AACTCCGCCG AGTTCGCCG ATTCTCCGG
1501 CATACGTTTC GTACGTAGAG TTAATCAGTC GTTGGTATCA GGGCGGGGAT TCGAGCGGGT AGGCGGGG TTAGAGCGGG TCAAGCGGG TAAGAGCGG
1601 CCATGGCTGA CTAATTTTTT TTAATTTATG AGAGCCGAG CCCTCTCTG CCTCTGAGCT ATTCCAGAG TAGTGAGGAG GCTTTTTGG AGCCTAGG
1701 GGTACCGACT GATTAAAGAA TCGTTTTCGA TGAATGATG AGATGATG CAGCAGTTT TCCCGCCG TCCGATCTTC TAAGGTTCTC CGAAAAAAC CCGATCCG
1801 AAAACGTTTT TCGAGCTCCT AGCAAAGCT ACTAATTTGT TCTACCTAAC GGTGCTCAA GAGCGCGGG CCGGTTCTT TTTGTCAA GCGACTGTG CCGTCCCTG
1901 GGCAACAACG ACAATCGGCT GCTCTGATG CGCGTGTTC TCGTGTGCTG CCAGCAAGT GCGTCCCTG GCGCAAGT TCGAGTTGT TCGAGTTGT GGAAGGACT
2001 TTAGTTGAC TCCTGCTCCG TCGCGCCGAT AGCACCGAC GGTGCTGCC ATCTCACCTT GCTCTGCC AGAAGTATC CATCATGGT GATGCAATG GCGGCTGCA
2101 GGCTGCTATT GGGCGAAGTG CCGGCGCAG ATCTCTGTC TAGAGTGA TCGAGCGAG ACCTACTCG ATGAAAGCG GTCTTCTAG CTACGTTAG CCGCGAGCT
2201 CCGAGATTA CCGGCTTAC GCGGCTTCC TAGAGTGA CCACCAAG CGGTAAGT TCGGTTGCT TCGGTTGCT GCGGTTGCT GCGGTTGCT GCGGTTGCT
2301 ATGCGAACA GTGCGATGGA GCGGTAAGT GCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT
2401 GACCTGTTTC TCGTAGTCCC GTGGAATCAT GTGGAATCAT GTGGAATCAT GTGGAATCAT GTGGAATCAT GTGGAATCAT GTGGAATCAT GTGGAATCAT
2501 TACCGTGAT ATTGCTGAAG AGCTTGGGG CGAATGGGCT GACGCTTCC TCGTCTTAA CGGTATCGG GCGTATCGG GCGTATCGG GCGTATCGG
2601 CGCTTCTTG TAACGACTTC TCGAACCGC GCTTACCGA GCGTATCGG GCGTATCGG GCGTATCGG GCGTATCGG GCGTATCGG GCGTATCGG
2701 AAGSCCAGCA AAGGCCAGG AACCGTAAA AGGCGCGT TCGGCGTTC TCGGCGTTC TCGGCGTTC TCGGCGTTC TCGGCGTTC TCGGCGTTC
2801 TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGTACCAG GCGTTCGCC TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT
2901 AGTCTCCAC GCTTTGGGT GTCTGTGAT TCTATGTT TCTATGTT TCTATGTT TCTATGTT TCTATGTT TCTATGTT TCTATGTT
3001 ATGACACAGG GGAAGAGGG AAGCCCTTCG CACGCGGAA GAGTATCGAG TCGGACTCC ATAGAGTCAA TATGCGAGCT GAGATCGAT TACACTCGT
3101 GTGTGCAGCA ACCCCCGT CAGCCCGAC CAGCCCGAC CAGCCCGAC CAGCCCGAC CAGCCCGAC CAGCCCGAC CAGCCCGAC CAGCCCGAC
3201 AGCCACTGTT TGGGGGCA GTCCGGCTGG CAGCGCGGAA GTATGTAGG GTGCTACAG AGTTCCTGAA TCGGTTGCT TCGGTTGCT TCGGTTGCT
3301 GGTATCTGG CTCTGTGAA GCGAGTTACC TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT
3401 CCATAGACGC GAGACGACTT CGTCAATAGT AAGCTTTT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT TCGGTTGCT
3501 AGCAGCAGT TACGCGCA AAAAAGGAT CTCGAAGA TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG
3601 TCGTCTGCTA ATGCGCGTCT TTTTCTCTA GAGTCTTCT AGGAATATG AAAATATG TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG
3701 TAATCAGTGA GGCACCTATC TCAGGATCT GTCTATTTC TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG TCGTTCATG
3801 ATTAGTCACT CCGTGGATAG AGTCCGCTAG CAGATAAGG AAGTATGAT CAGGACTG CAGGACTG CAGGACTG CAGGACTG CAGGACTG
3901 ATCTGGCCCC AGTCTGCAA TGATACCGG AGACCCAGG TCACGGGCT CAGGTTATC AGCAATAAG AGGAGGAG CAGGAGGAG CAGGAGGAG
4001 TAGACCGGG TCACGACGTT ACTATGGCG AGTGGGCTG AGTGGGCTG AGTGGGCTG AGTGGGCTG AGTGGGCTG AGTGGGCTG AGTGGGCTG
4101 GGTCCTGCAA CTTTATCCG CTCCATCCAG TCTATTAAT GTTGCCGGGA AGTAGAGTA AGTAGAGTA AGTAGAGTA AGTAGAGTA

FIGURE 5 CONT.

9/11

3401 CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT TCGATCTCAT TCATCAAGCG GTCAATTATC AAACGCGTTG CAACAACGGT
TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTG AGCTCCGGTT CCCAACGATC AAGCGAGTT ACATGATCCC CCATGTTGTG
AACGATGTCC GTAGCACAC AGTGGAGCA CCAACCATTA CCGAAGTAAG TCGAGGCCAA GGGTTGCTAG TCCGCTCAA TGTAAGGG GTACAACAC
CAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTGTG GATCGTGTG AGAAGTAAGT TGGCGCAGT GTTATCACTC ATGTTATGG CAGCACTGCA TAATTCTCTT
3501 GTTTTTCGC CAATCGAGGA AGCCAGGAG CTAGCAACAG TCTTCATTCA ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA
ACTGTCAATG CATCCGTAAG ATGCTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTG TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG
3601 TGACAGTACG GTAGGCATTC TACGAAAGA CACTGACCAC TCAATGATTG GTTCAGTAAG ACTCTTATCA CATACGCCCG TGGCTCAACG AGAACGGGC
CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGCGGAAA CTCTCAAGGA TCTTACCGCT
3701 GCAGTTATGC CCTATTATGG CGCGTGATAT CGTCTGAAA TTTTCACGAG TAGTAACCTT TTGCAAGAAG CCCCCTTTT GAGAGTTCCT AGAATGGCGA
GTTGAGATCC AGTTCGATGT AACCCACTCG TGCACCCAAC TGATCTTCAG CATCTTTTAC TTTCAACCAGC GTTCTGGGT GAGCAAAAAC AGGAAGGCAA
3801 CAACTCTAGG TCAAGCTACA TTGGTGAGC ACGTGGGTTG ACTAGAAATG GTAGAAAATG AAAGTGCTG CAAAGACCCA CTCGTTTTTG TCCTTCCGTT
AATGCCGCAA AAAAGGGAAT AAGGGGACA CGGAATGTT GAATACTCAT ACTCTTCTT TTTCAATATT ATTGAAGCAT TTATCTAGAG GTTATTGTCT
3901 TTACGGCGTT TTTTCCCTTA TTCCGCTGT GCCTTACAA CTTATGAGTA TGAGAAGGA AAGTTATAA TNACTTCGTA ANTAGATCTC CAATAACAGA

FIGURE 5 CONT.



1 AGTGTCTAG ACCTGTTGAC AATTAATCAT CGGCTCGTAT AATGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGGATCG ATCGAATTTCG
101 TCACGAGATC TGGACAACTG TTAATTAGTA GCCGAGCATA TPACACACCT TAACACTCGC CTATTGTAA AGTGTGTCTT TTGTCTTAGC TAGCTTAAGC
GATCCAAGCT TGAGCTCGAG CCATGGCCCG GGTACCTAAT TAGAARAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGGTAATC TGCTGCTTGC
201 CTAGTTTCTGA ACTCGAGCTC GGTACCGGGC CCACTTAATA AGCTTTTCTA GTTTCCTAGA AGAACTCTAG GAAAAAAGA CGGCAATTAG ACGACGAAACG
AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTGTGTTGCC GCATCAAGAG CTACCAACTC TTTTCCGAA GGTAACTGGC TTCCAGCAGAG CGCAGATACC
301 TTTGTTTTTT TGGTGGCGAT GGTGCGCACC AAACAACGG CCTAGTTCTC GATGTTGAG AAAAAGGCTT CCATTGACCG AAGTCGTCTC GCGTCTATGG
AAATACTGTT CTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
401 TTTATGACAA GAAGATCACA TCGGCATCAA TCCGGTGGTG AAGTTCTTGA GACATCTGG CGATGTATG GAGCGAGACG ATTAGGACAA TGGTCACCGA
GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGTGTGACT CAAGACGATA GTTACCGGAT AAGCGCGAGC GGTGCGGCTG AACGGGGGT TCGTGCACAC
501 CGACGGTCAC CGCTATTTCAG CACAGAAATGG CCCAACTGA GTTCTGCTAT CAATGGCTTA TTCCGCGTCA TCCGCGCTCG CCAGCCCCGAC TTGCCCCCA ACGACGTGTG
AGCCCAAGCTT GGAGCGAACG ACCTACACCG AACTAGATA CCTACAGCGT GAGCTATGAG AAAGCGCAC GGTTCGCCAA GGTTCGCCA CGCAGGTGTG
601 TCGGTCGAA CCTCGCTTGC TGGATGTGGC TTGACTCTAT TGATGCTCA GCGCACGAG GAGCTTCCAG GGGGAACGC CTGGTAICTT TATAGTCTG TCGGGTTTCG CCACCTCTGA
TCCGTAAGC GGCAGGGTCG GAACAGAGA CCGTGTCTC CGCTGCTCC CTCGAAGGTC CCCCTTTGCG GACCATAGAA ATATCAGAC AGCCAAAGC GGTGAGACT
701 AGGCCATTTC GATTTTGTG ATGCTCTCA GGGGGGCGA GCCTATGAA AACGCCAGC AACCGGCTT TTTTACGGTT CTGCCCCGT CATTAGCGCG
GAACCTGCGC CTAAACAC TACGAGCAG CTATCTCAGC GATCTGCTA TTTCGTTTAT CCATAGTGC CTGACTCCCC GTGCTGTAGA TAACTACGAT
801 GCTATTACCA ATGCTTAATC AGTAGGACAT TCACCTCGTG GATAGAGTCG CTAGACAGAT AAAGCAAGTA GGTATCGAG GACTGAGGG CAGCACATCT AATTGATGCTA
ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTACC GGCTCCAGAT TTATCAGCAA TAAACACAGC AGCGGGAAGG
901 TGCCCTCCCG AATGTFAGAC CGGGGTCAAC ACGTTACTAT GCGCTCTGG GTGCGAGTGG CCGAGGTCTA AATAGTCTGT ATTTGGTCTG TCGGCCCTTCC

FIGURE 6

11/11

1001 GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC
CGGCTCGCGT CTTCAACAGG ACGTTGAAAT AGGCGGAGGT AGGTGAGATA ATTAACAACG GCCCTTCGAT CTCATTTCATC AAGCGGTCAA TTATCAAACG
1101 GCAACGTTGT TGCCATTGCT ACAGGCATCG TGGTGTCACG CTCGTGCTTT GGTATGGCTT CAATCAGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG
CGTTGCAACA ACGGTAACGA TGTCGTAGC ACCACAGTGC GAGCAGCAAA CCAATCCGAA GTAAGTCCGAG GCCAAGGGTT GCTAGTCCG CTCAATGTAC
1201 ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG TTGTCAGAACG GCACTGTTAT CACTCATGGT TATGCGAGCA
TAGGGGTAC AACACGTTTT TTCGCCAATC GAGGAAGCCA GGAGGTAGC AACAGTCTTC ATTCAACCGG CGTCACAATA GTGAGTACCA ATACCGTCGT
1301 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC ACCACTCATG TCAACCAAGT CAATCTGAGA ATAGTGATG CGGCGACCGA
GACGTATTAA GAGAATGACA GTACGGTAGG CAATCTACGA AAAGACACTG AACTTAAAG TGCTCATCAT GTAAGACTCT TATCACATAC GCCGTGGCT
1401 GTTGCTTTG CCCGCGTCA ATACGGGATA ATACCGGCC ACATAGCAGA ACTTTAAAG TGCTCATCAT TGGAAAACGT TCTTCGGGC GAAAACCTCTC
CAACAGAAC GGGCGCGAGT TATGCCCTAT TATGCGCGG TGTATCGTCT TGAATTTTC ACGAGTAGTA ACCTTTTCGA AGAAGCCCCG CTTTGTGAGAG
1501 AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC ACTCGTGCTC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA
TTCTAGAAAT GCGGACAACT CTAGGTCAAG CTACATGGG TGAGCACGAG GGTGACTAG AAGTCGTAGA AAATGAAAAG GGTGCAAG ACCCACTCGT
1601 AAAACAGGAA GGCAAAATGC CGCAAAAAG GGAATAAGG CGACAGGAA ATGTTGAATA CTCATCTCT CTCTTTTCA ATATTATTGA AGCATTATC
TTTTTCCTT CCGTTTACG GCGTTTTC CCTTATCCC GCTGTGCCCT TACAACCTTAT GAGTATGAGA AGGAAAAGT TATAATACT TCGTAAATAG
1701 AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT CTAGAGGTA
TCCCAATAAC AGAGTACTCG CCTATGTATA AACTTACATA GATCTTCCAT

FIGURE 6 CONT.